Supplementary Materials

Supplementary Materials and Methods

Genome sequencing, assembly, and annotation

DNA was extracted from the opah muscle tissues for genome sequencing. Nanopore reads were produced using the Oxford Nanopore GridIon platform. The BGI read libraries were constructed using a TruSeq DNA PCR-Free Library Prep Kit and sequenced using the BGI MGISEQ-2000 platform. A total of 152.97 Gb of Nanopore reads and 102.03 Gb of BGI reads were generated. The BGI reads were used to estimate genome size with the k-mer method. The 17-mer depth frequency distribution was calculated using Jellyfish (Marçais & Kingsford, 2011). The predicted genome size was 1 175 Mb (**Supplementary Figure S4**). The wtdbg2 v2.4.1 program (Ruan & Li, 2020) was used to assemble the contigs. The BGI reads were mapped to the genome using BWA-MEM (Li, 2013) with default parameters. Two rounds of polishing were then performed using NextPolish v1.0 (Hu et al., 2020) based on aligned files. BUSCO (Simão et al., 2015) was used to assess genome quality and gene completeness with the library "actinopterygii_odb9". Gene annotation was conducted using a combination of *ab initio* and homology-based methods. The *ab initio* approach was conducted using Augustus v3.2.1 (Stanke et al., 2008), SNAP (Leskovec & Sosic, 2017), and GlimmerHMM (Majoros et al., 2004). For homology-based annotation, the protein sequences of *Danio rerio*, *Oryzias latipes*, and *Gadus morhua* were downloaded from Ensembl (http://www.ensembl.org/index.html) and aligned to the opah genome using tBLASTn (E-value≤1e-05) (Altschul et al., 1990). GeneWise v2.4.1 (Birney et al., 2004) was then used to identify accurate gene structures.

We extracted mitochondrial genes *COI* and cyt *b* to identify species. An EZNA® Tissue DNA Kit (OMEGA, Wuhan, China) was employed to extract DNA from opah muscle. The mitochondrial *COI* gene was amplified by polymerase chain reaction (PCR) using primers FishF1-COI 5' TCA ACC AAC CAC AAA GAC ATT GGC AC and FishR1-COI 5' TAG ACT TCT GGG TGG CCA AAG AAT CA. The mitochondrial cyt *b* gene was amplified by PCR using primers L14504-ND6 5' GCC AAW GCT GCW GAA TAM GCA AAG GTG and H15149-CYB 5' GCK CCT CAG AAG GAC ATT TGK CCT CA. Previously published opah *COI* and cyt *b* genes were downloaded from GenBank (accession# JF931865-JF931967) (Hyde et al., 2014) and aligned with the counterpart genes of opah from this study. Each gene set was aligned using MUSCLE and a maximum-likelihood tree was constructed in MEGA v7 (Kumar et al., 2016) with default parameters.

Phylogenetic analysis

The protein sequences of zebrafish (*Danio rerio*; accession: GRCz11), medaka (*Oryzias latipes*; accession: ASM223469v1), Atlantic cod (*Gadus morhua*; accession: gadMor3.0), platyfish (*Xiphophorus maculatus*; accession: X_maculatus-5.0-male), stickleback (*Gasterosteus aculeatus*; accession: BROAD S1), Mexican tetra (*Astyanax mexicanus*; accession: Astyanax_mexicanus-2.0), Japanese puffer (*Takifugu rubripes*; accession: fTakRub1.2), spotted gar (*Lepisosteus oculatus*; accession: Eluc_v4), Amazon molly (*Poecilia formosa*; accession: Poecilia_formosa-5.1.2), climbing perch (*Anabas testudineus*; accession:fAnaTes1.2), Nile tilapia (*Oreochromis niloticus*; accession: O_niloticus_UMD_NMBU), big-finned mudskipper (*Periophthalmus magnuspinnatus*; accession: PM.fa), and northern pike (*Esox Lucius*; accession: Eluc_v4) were downloaded from Ensembl. The protein sequences were filtered based on the longest transcripts and low-quality sequences. Orthologous groups were constructed by OrthoFinder (Emms & Kelly, 2015) using the filtered sequences. A total of 2 441 one-to-one orthologous genes in opah and 13 ectothermic fish (*Danio rerio*, *Oryzias latipes*, *Gadus morhua*, *Xiphophorus maculatus*, *Gasterosteus aculeatus*, *Astyanax mexicanus*, *Takifugu rubripes*, *Lepisosteus oculatus*, *Poecilia formosa*, *Anabas testudineus*, *Oreochromis niloticus*, *Periophthalmus*

magnuspinnatus, and *Esox lucius*) were identified. Multi-sequence alignments were generated at the protein level for the one-to-one orthologous genes using MAFFT v7.453 (Katoh & Standley, 2013). Gblocks v0.91b (Castresana, 2000) was used to extract the conserved sites in the multi-sequence alignment. Based on the filtered multi-sequence alignment, we constructed a phylogenetic tree using IQ-tree v2.0.3 with default parameters (Lam-Tung et al., 2015). MCMCTREE (PAML v9.1i) (Yang, 2007) was used to estimate opah divergence times. Fossil calibration information (divergence time of *Oryzias latipes* and *Xiphophorus maculatus* is 79–107 million years, *Xiphophorus maculatus* and *Oreochromis niloticus* is 88–139 million years, *Gasterosteus aculeatus* and *Takifugu rubripes* is 99–127 million years, *Danio rerio* and *Lepisosteus oculatus* is 295–334 million years, and *Danio rerio* and *Lampris megalopsis* is 206–252 million years) was obtained from http://timetree.org. The end time node used was 315 million years (http://timetree.org/). The gamma prior for the rate for genes was set to 1, 3.565 based on substitution rates inferred using BaseML. Tracer (Andrew et al., 2018) was used to analyze the MCMC.txt file and detect ESS (Effective Sample Size) values. The ESS values were all higher than 200, indicating convergence.

Positive selection and rapid evolution analysis

The oarfish (*Regalecus glesne*; accession: GCA_900302585.1) is the closest species to opah (Malmstrøm et al., 2016), and thus we downloaded genomic data of the oarfish from the NCBI database (https://www.ncbi.nlm.nih.gov/). The corresponding oarfish genome sequence was searched against the single-copy orthologous proteins of opah using aTRAM v2.0 (Allen et al., 2018). Orthologous groups were constructed using OrthoFinder v.2.2.7 (Emms & Kelly, 2015) with filtered sequences of six fish species (i.e., opah, oarfish, stickleback, medaka, Atlantic cod, and northern pike). In total, 6 119 one-to-one orthologous genes from five species (oarfish, stickleback, medaka, Atlantic cod, and northern pike) and opah were identified and compared. All orthogroups of the six species were aligned based on the coding sequences using Prank v.170427 (Löytynoja, 2014). Alignments gaps and ambiguous positions were removed using Gblocks under parameters -t=c -b5=h (Castresana, 2000). The base-substitution mutation rate of non-synonymous mutations is *Ka* and the base-substitution mutation rate of synonymous mutations is *Ks*. *Ka*/*Ks*>1 (i.e., positive selection) indicates that most non-synonymous mutations are favorable and are evolutionarily retained, and thus positive selection is the main driving force of biological evolution. To detect positively selected genes, the branch-site model applied in CODEML (PAML v9.1i) (Yang, 2007) was used, with opah used as the foreground branch. *P*-values were calculated by likelihood ratio tests using a chi-square distribution. Furthermore, the false discovery rate (FDR) function in the R package was applied to correct *P*-values. Positive sites with an FDR-corrected *P*-value>0.05 and no * significant marker (higher than 0.95) for Bayes Empirical Bayes (BEB) were filtered out. For the remaining positively selected genes, MEGA (Kumar et al., 2016) was used to visually check whether the positive selection sites were located in poorly aligned regions. Thus, 158 genes were identified as under positive selection.

To identify genes under rapid evolution, we calculated the *Ka/Ks* values for single-copy genes using the branch model in CODEML, with opah as the foreground branch. *P*-values were calculated as described above. A higher *Ka/Ks* ratio on the foreground branch is classified as under rapid evolution, i.e., genetic change occurs rapidly enough to have a measurable impact on simultaneous ecological change. In total, 44 genes were identified as under rapid evolution.

Gene family expansion and contraction analysis

A gene family is a set of several similar genes, formed by duplication of a single original gene and generally with similar biochemical functions. Because of the low quality of the oarfish genome, gene family expansion and contraction analysis was performed for five

species (opah, stickleback, medaka, Atlantic cod and northern pike) using Café v3.1 (Han et al., 2013). Analysis indicated that 140 gene families were expanded and 52 gene families were contracted.

Identification of candidate genes involved in endothermy

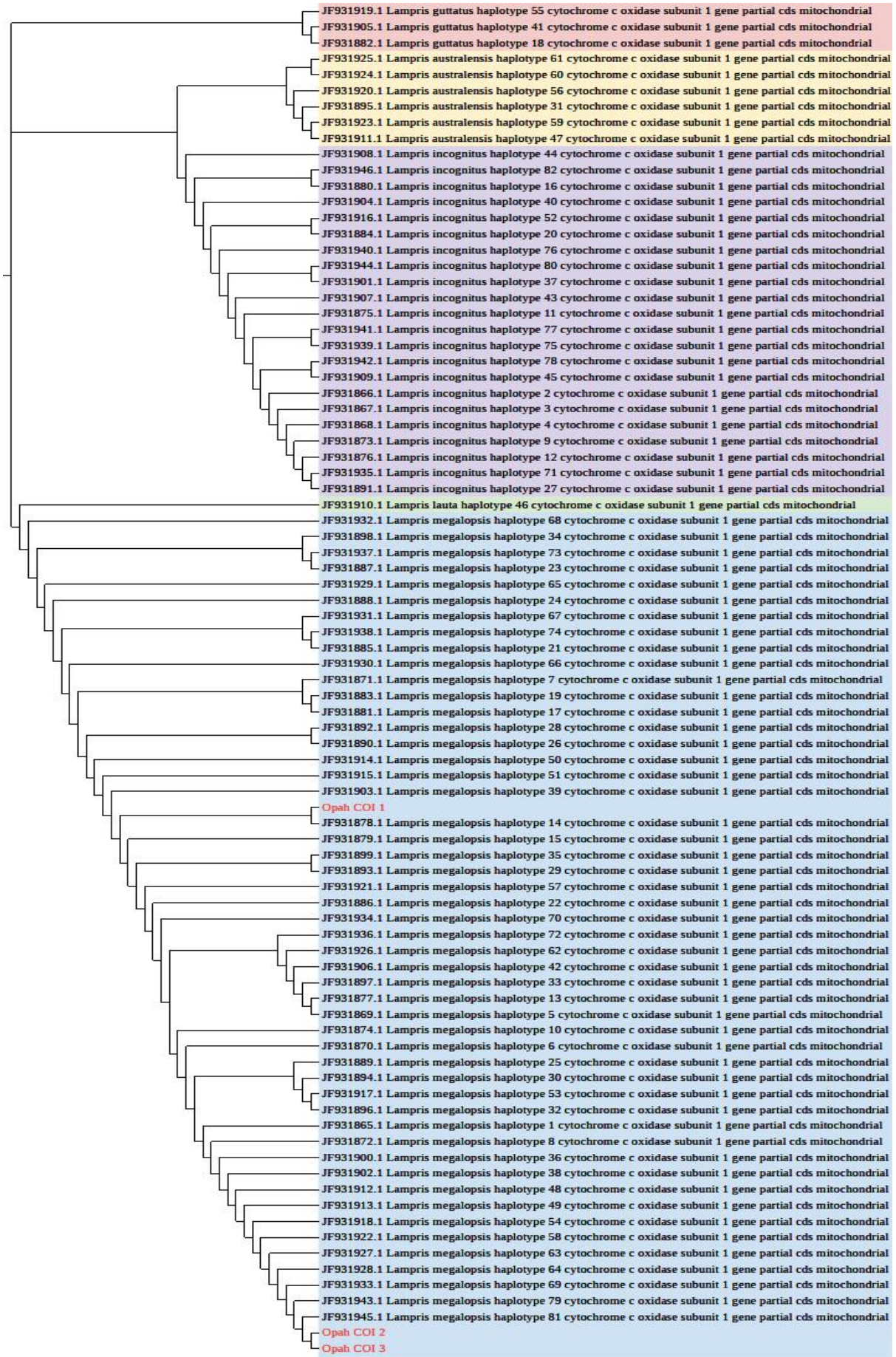
Positively selected genes, rapidly evolving genes, and expanded gene families were annotated using Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.genome.jp/kaas-bin/kaas_main). Eggnog (Huerta-Cepas et al., 2018) annotation and InterPro (Hunter et al., 2012) structural annotation results were integrated into Gene Ontology (GO) annotation, with GO enrichment analysis then performed using Ontologizer (Bauer et al. 2008). According to thermogenesis and heat preservation, the enriched candidate genes were divided into muscle differentiation and development, enhancement of muscle contractile function, metabolic energy supply, reduction of lactic acid accumulation, *Sln*-mediated thermogenesis, formation of retia mirabilia, and fat formation and metabolism. SMART (Letunic et al., 2011) was used to view gene domain information. We also applied PolyPhen-2 (Adzhubei et al. 2010) and PROVEAN (Choi & Chan, 2015) to analyze the effects of the mutation sites on proteins, and Phyre2 (Kelley et al., 2015) to predict the 3D structure.

***Ldha* expression and enzyme activity measurement**

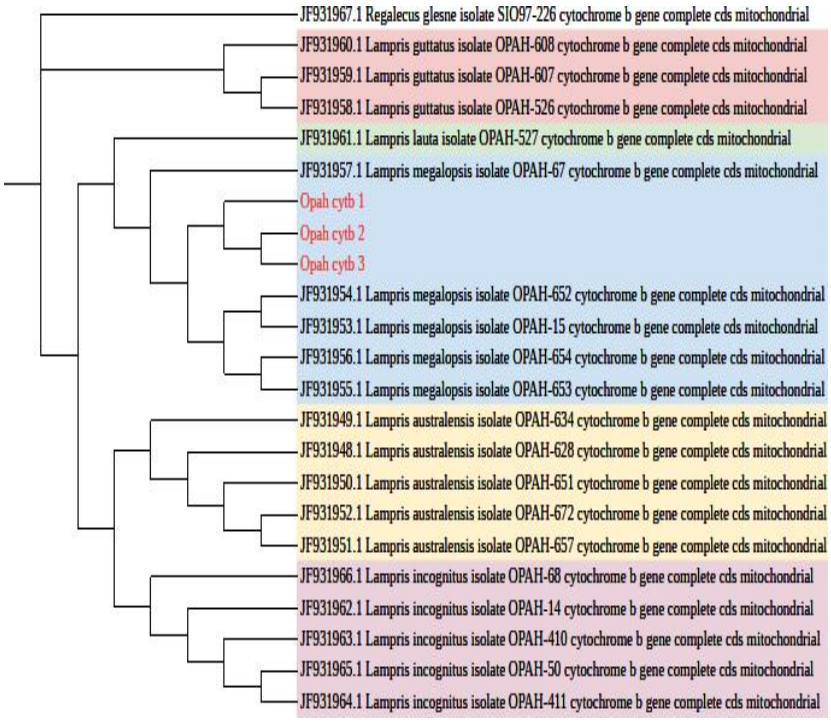
High activities of both lactate dehydrogenase (LDH) and citrate synthase (CS) in lamnid sharks compared to their ectothermic relatives (Bernal et al., 2003). At present, however, there is no relevant research on opah. Analysis showed that *Ldha* was a positively selected gene in opah. Therefore, we tested whether the positively selected sites improved LDH efficiency in opah and compared LDH enzyme activity in opah and stickleback fish.

Genes were synthesized by the Frdbio Company (Wuhan, China), including *Ldha*-opah, *Ldha*-stickleback, and *Ldha2*-opah-replace (C261V and C271V). All genes were codon-optimized based on *Escherichia coli*. Three expression vectors (Frd-HIS-opah, Frd-HIS-stickleback, and Frd-HIS-261V-271V) were constructed and transfected into *E. coli* BL21 (DE3) cells in the presence of 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). We then analyzed expression by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and purified the target protein by Sephadex. An LHD assay kit (Lai Er Bio-Tech, Hefei, China) was used to measure enzyme activity according to the manufacturer's instructions. We prepared a standard curve, and tested samples and controls. Optical density (OD) was measured at 450 nm, and enzyme activity was calculated based on the absorbance and standard curve. A *t*-test was used to analyze the data.

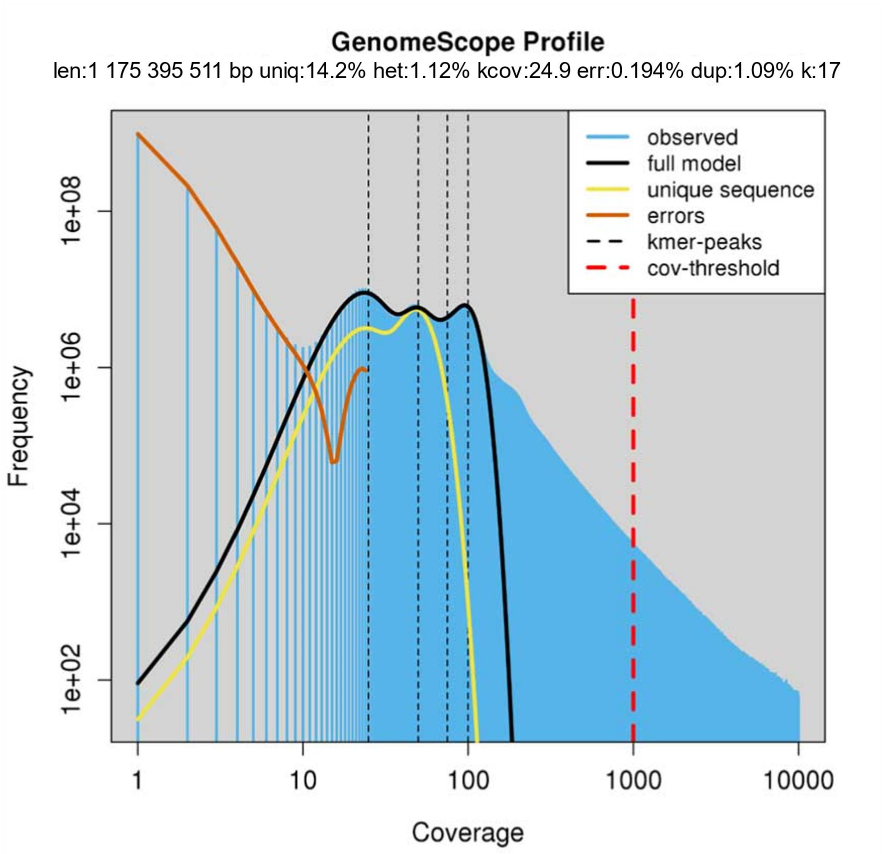
Supplementary Figure S1 Sample collection. Picture on left was taken onboard “PING TAI RONG 65” during opah collection. Right two pictures show Zhoushan dissection of opah muscle tissue.


JF931919.1 Lampris guttatus haplotype 55 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931905.1 Lampris guttatus haplotype 41 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931882.1 Lampris guttatus haplotype 18 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931925.1 Lampris australensis haplotype 61 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931924.1 Lampris australensis haplotype 60 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931920.1 Lampris australensis haplotype 56 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931895.1 Lampris australensis haplotype 31 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931923.1 Lampris australensis haplotype 59 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931911.1 Lampris australensis haplotype 47 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931908.1 Lampris incognitus haplotype 44 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931946.1 Lampris incognitus haplotype 82 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931880.1 Lampris incognitus haplotype 16 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931904.1 Lampris incognitus haplotype 40 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931916.1 Lampris incognitus haplotype 52 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931884.1 Lampris incognitus haplotype 20 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931940.1 Lampris incognitus haplotype 76 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931944.1 Lampris incognitus haplotype 80 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931901.1 Lampris incognitus haplotype 37 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931907.1 Lampris incognitus haplotype 43 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931875.1 Lampris incognitus haplotype 11 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931941.1 Lampris incognitus haplotype 77 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931939.1 Lampris incognitus haplotype 75 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931942.1 Lampris incognitus haplotype 78 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931909.1 Lampris incognitus haplotype 45 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931866.1 Lampris incognitus haplotype 2 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931867.1 Lampris incognitus haplotype 3 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931868.1 Lampris incognitus haplotype 4 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931873.1 Lampris incognitus haplotype 9 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931876.1 Lampris incognitus haplotype 12 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931935.1 Lampris incognitus haplotype 71 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931891.1 Lampris incognitus haplotype 27 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931910.1 Lampris lauta haplotype 46 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931932.1 Lampris megalopsis haplotype 68 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931898.1 Lampris megalopsis haplotype 34 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931937.1 Lampris megalopsis haplotype 73 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931887.1 Lampris megalopsis haplotype 23 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931929.1 Lampris megalopsis haplotype 65 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931888.1 Lampris megalopsis haplotype 24 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931931.1 Lampris megalopsis haplotype 67 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931938.1 Lampris megalopsis haplotype 74 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931885.1 Lampris megalopsis haplotype 21 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931930.1 Lampris megalopsis haplotype 66 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931871.1 Lampris megalopsis haplotype 7 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931883.1 Lampris megalopsis haplotype 19 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931881.1 Lampris megalopsis haplotype 17 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931892.1 Lampris megalopsis haplotype 28 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931890.1 Lampris megalopsis haplotype 26 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931914.1 Lampris megalopsis haplotype 50 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931915.1 Lampris megalopsis haplotype 51 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931903.1 Lampris megalopsis haplotype 39 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
Opah COI 1
JF931878.1 Lampris megalopsis haplotype 14 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931879.1 Lampris megalopsis haplotype 15 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931899.1 Lampris megalopsis haplotype 35 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931893.1 Lampris megalopsis haplotype 29 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931921.1 Lampris megalopsis haplotype 57 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931886.1 Lampris megalopsis haplotype 22 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931934.1 Lampris megalopsis haplotype 70 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931936.1 Lampris megalopsis haplotype 72 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931926.1 Lampris megalopsis haplotype 62 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931906.1 Lampris megalopsis haplotype 42 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931897.1 Lampris megalopsis haplotype 33 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931877.1 Lampris megalopsis haplotype 13 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931869.1 Lampris megalopsis haplotype 5 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931874.1 Lampris megalopsis haplotype 10 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931870.1 Lampris megalopsis haplotype 6 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931889.1 Lampris megalopsis haplotype 25 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931894.1 Lampris megalopsis haplotype 30 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931917.1 Lampris megalopsis haplotype 53 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931896.1 Lampris megalopsis haplotype 32 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931865.1 Lampris megalopsis haplotype 1 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931872.1 Lampris megalopsis haplotype 8 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931900.1 Lampris megalopsis haplotype 36 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931902.1 Lampris megalopsis haplotype 38 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931912.1 Lampris megalopsis haplotype 48 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931913.1 Lampris megalopsis haplotype 49 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931918.1 Lampris megalopsis haplotype 54 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931922.1 Lampris megalopsis haplotype 58 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931927.1 Lampris megalopsis haplotype 63 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931928.1 Lampris megalopsis haplotype 64 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931933.1 Lampris megalopsis haplotype 69 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931943.1 Lampris megalopsis haplotype 79 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
JF931945.1 Lampris megalopsis haplotype 81 cytochrome c oxidase subunit 1 gene partial cds mitochondrial
Opah COI 2
Opah COI 3

Supplementary Figure S2 Previously published opah *COI* genes (GenBank, accession# JF931865-JF931947) (Hyde et al., 2014) were aligned with counterpart genes of opah from this study. Genes were aligned using MUSCLE and a maximum-likelihood tree was constructed using MEGA v7 (Kumar et al., 2016) with default parameters. Opah from this study was clustered with *COI* set of *L. megalopsis*.



Supplementary Figure S3 Previously published opah cyt *b* genes (GenBank, accession# JF931948-JF931967) (Hyde et al., 2014) were aligned with counterpart genes of opah from this study. Genes were aligned using MUSCLE and a maximum-likelihood tree was constructed using MEGA v7 with default parameters (Kumar et al., 2016). Opah from this study was clustered with cyt *b* set of *L. megalopsis*.



Supplementary Figure S4 Illumina reads were used to estimate genome size with Jellyfish (Marçais & Kingsford, 2011). 17-mer depth frequency distribution was calculated using k-mer method. Predicted genome size was 1 175 Mb

REFERENCES

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. 2010. A method and server for predicting damaging missense mutations. *Nature Methods*, **7**(4): 248–249.

Allen JM, LaFrance R, Folk RA, Johnson KP, Guralnick RP. 2018. aTRAM 2.0: An improved, flexible locus assembler for NGS data. *Evolutionary Bioinformatics*, **14**: 1176934318774546.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, **215**(3): 403–410.

Bauer S, Grossmann S, Vingron M, Robinson PN. 2008. Ontologizer 2.0—a multifunctional tool for GO term enrichment analysis and data exploration. *Bioinformatics*, **24**(14): 1650–1651.

Bernal D, Smith D, Lopez G, Weitz D, Grimminger T, Dickson K, et al. 2003. Comparative studies of high performance swimming in sharks II. Metabolic biochemistry of locomotor and myocardial muscle in endothermic and ectothermic sharks. *Journal of Experimental Biology*, **206**:2845-2857.

Birney E, Clamp M, Durbin R. 2004. GeneWise and genomewise. *Genome Research*, **14**(5): 988–995.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, **17**(4): 540–552.

Choi Y, Chan AP. 2015. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, **31**(16): 2745–2747.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, **16**(1): 157.

Han MV, Thomas GWC, Lugo-Martinez J, Hahn MW. 2013. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Molecular Biology and Evolution*, **30**(8): 1987–1997.

Hu J, Fan JP, Sun ZY, Liu SL. 2020. NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics*, **36**(7): 2253–2255.

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. 2018. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, **47**(D1): D309–D314.

Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, et al. 2012. InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Research*, **40**(D1): D306–D312.

Hyde JR, Underkoffler KE, Sundberg MA. 2014. DNA barcoding provides support for a cryptic species complex within the globally distributed and fishery important opah (*Lampris guttatus*). *Molecular Ecology Resources*, **14**(6): 1239–1247.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**(4): 772–780.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, **10**(6): 845–858.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**(7): 1870–1874.

Leskovec J, Sosic R. 2017. SNAP: a general-purpose network analysis and graph-mining library. *ACM Transactions on Intelligent Systems and Technology*, **8**(1): 1–20.

Letunic I, Doerks T, Bork P. 2011. SMART 7: recent updates to the protein domain annotation

resource. *Nucleic Acids Research*, **40**(D1): D302–D305.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv: 1303.3997v1.

Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. *In*. Russell DJ. Multiple Sequence Alignment Methods. Totowa: Humana Press, 155–170.

Majoros WH, Pertea M, Salzberg SL. 2004. TigrScan and GlimmerHMM: two open source *ab initio* eukaryotic gene-finders. *Bioinformatics*, **20**(16): 2878–2879.

Malmstrøm M, Matschiner M, Tørresen OK, Star B, Snipen LG, Hansen TF, et al. 2016. Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics*, **48**:1204-1210.

Marçais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. *Bioinformatics*, **27**(6): 764–770.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**(1): 268–274.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, **67**(5): 901–904.

Ruan J, Li H. 2020. Fast and accurate long-read assembly with wtdbg2. *Nature Methods*, **17**(2): 155–158.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, **31**(19): 3210–3212.

Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics*, **24**(5): 637–644.

Yang ZH. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**(8): 1586–1591.

Supplementary Table S1 Opah genome assembly

Term	Number
Total length	1090863164(base)
Number of contigs	12081
Number>=2000bp	12079
N50	590575
N90	30919
GC rate	0.383

Supplementary Table S2 Positively selected genes

PAML						
ID	Gene	MA0 lnL	MA lnL	*P*-value	FDR	Description
yueyu04924	Gli1	-14451.4646	-14446.75588	0.002149325	0.045784393	zinc finger protein Gli1，ko:K16797，Hedgehog signaling pathway ko04340；reference (Cheng et al.，2010)；
yueyu09933	Dcn	-3822.494326	-3816.442991	0.000503497	0.022475958	decorin，ko:K04660，TGF-beta signaling pathway ko04350；reference (Cheng et al.，2010)；
yueyu20061	Paxbp1	-9262.973401	-9254.996039	0.000064876	0.005967609	PAX3- and PAX7- binding protein 1；GO:0014857 regulation of skeletal muscle cell proliferation；
yueyu08135	Unc45a	-10825.63979	-10820.6228	0.001536778	0.038320788	protein unc-45 homolog A ；GO:0030239 myofibril assembly;
yueyu02147	Unc45b	-11223.87937	-11216.10839	0.000080694	0.006620179	protein unc-45 homolog B ；GO:0030239 myofibril assembly;
yueyu01092	Tmod4	-3950.716124	-3945.27112	0.000966844	0.030290821	tropomodulin-4;GO:0030239 myofibril assembly
yueyu23591	Myom2	-17051.73007	-17044.79322	0.000195517	0.012241827	myomesin-2; GO:0030239 myofibril assembly
yueyu19087	Myom3	-10206.75232	-10201.83626	0.001714902	0.03973729	myomesin-3; GO:0030239 myofibril assembly
yueyu15403	Tcap	-2500.905762	-2496.256825	0.002294199	0.047213838	telethonin; GO:0030239 myofibril assembly
yueyu01534	M-protein	-16689.31566	-16677.74545	0.000001506	0.0000361	M-protein, striated muscle; GO:0006941 striated muscle contraction;
yueyu03865	Pdgfra	-9954.425822	-9948.577637	0.000626221	0.025687755	ko:K04363, platelet-derived growth factor receptor alpha; Calcium signaling

						pathway ko04020;
yueyu11225	Plcb	-11966.33447	-11960.99462	0.001083173	0.032235016	ko:K05858, phosphatidylinositol phospholipase C, beta; Calcium signaling pathway ko04020; Adrenergic signaling in cardiomyocytes ko04261;
yueyu09243	Phka_b	-13553.30066	-13548.77835	0.002634678	0.049819032	ko:K07190, phosphorylase kinase alpha/beta subunit; Calcium signaling pathway ko04020;
yueyu16029	Casq2	-4755.048474	-4748.386611	0.000262069	0.014361944	ko:K23445, calsequestrin 2; Calcium signaling pathway ko04020; Cardiac muscle contraction ko04260;
yueyu10204	Ldha	-3661.964172	-3657.195994	0.002014413	0.043521357	ko:K00016,L-lactate dehydrogenase;Glycolysis / Gluconeogenesis ko:00010;
yueyu04057	Dlat	-6753.411267	-6748.352391	0.00146848	0.038262412	ko:K00627,pyruvate dehydrogenase E2 component (dihydrolipoamide acetyltransferase);Glycolysis / Gluconeogenesis ko:00010; Citrate cycle (TCA cycle) ko:00020
yueyu12498	Pgm2	-6994.220413	-6990.825995	0.00917297	0.048922507	ko:K15779,phosphoglucomutase / phosphopentomutase; Glycolysis /Gluconeogenesis ko:00010;
yueyu19957	Ryr1a	-48398.14257	-48393.59461	0.002561803	N/A	
yueyu09933	Dcn	-3822.494326	-3816.442991	0.000503497	0.022475958	decorin; GO:0038084 vascular endothelial growth factor signaling pathway;
yueyu20569	Aact	-4373.093415	-4365.196512	0.000070633	0.006127894	ko:K00626, acetyl-CoA

						C-acetyltransferase; Fatty acid degradation ko00071;

Supplementary Table S3 Rapidly evolving genes

ID	Gene	Foreground-branch w	Background-branch w	*P*-value	FDR	Description
yueyu06969	Mef2	0.06059	0.02921	0.002726291	0.021415316	ko:K04454,MADS-box transcription enhancer factor 2C;ko04371 Apelin signaling pathway
yueyu08607	Insig1	0.07922	0.03345	0.005914983	0.036106754	insulin-induced gene 1 protein;GO:0045444 fat cell differentiation; GO:0033993response to lipid; GO:0006633 fatty acid biosynthetic process;
yueyu04905	Ebf2	0.06192	0.02175	0.0000035	0.000191459	transcription factor COE2 isoform X1;GO:0045444 fat cell differentiation
yueyu14380	Adipor	0.10914	0.01323	0.003080246	0.023351487	ko:K07297,adiponectin receptor; ko04152 AMPK signaling pathway;ko04920 Adipocytokine signaling pathway
yueyu00168	Serca	0.05877	0.03621	0.009429217		serca1;

Supplementary Table S4 Expanded genes

ID	Gene	Divergence_size	Species_size	*P*-value	Description
yueyu06110 yueyu05829 yueyu05830	Htr2	1	3	0.036	ko:K04157, 5-hydroxytryptamine receptor 2; Calcium signaling pathway ko04020;
yueyu09028 yueyu00894 yueyu10159 yueyu10839 yueyu13489 yueyu14928 yueyu16454 yueyu16455	Camk2	5	8	0.021	ko:K04515, calcium/calmodulin-dependent protein kinase (CaM kinase) II; Calcium signaling pathway ko04020;
yueyu05372 yueyu05373 yueyu05374 yueyu05375 yueyu05376 yueyu15050 yueyu18181 yueyu00846 yueyu01023 yueyu06276 yueyu06280 yueyu06284 yueyu09664 yueyu09665	Igh	9	14	0	ko:K06856, immunoglobulin heavy chain;Calcium signaling pathway ko04020;

yueyu01263 yueyu17943	Ang-1	1	2	N/A	angiopoietin-1;reference (Kidoya H, 2015)